AIRCRAFT BASED REAL TIME BUNDLE ADJUSTMENT AND DIGITAL SURFACE MODEL GENERATION

P. d'Angelo and F. Kurz

German Aerospace Center (DLR), Remote Sensing Technology Institute, D-82234 Wessling, Germany
email: {Pablo.Angelo, Franz.Kurz}@dlr.de

Commission II, WG II/2

KEY WORDS: image orientation, real-time, dense matching, airborne camera, digital elevation model

ABSTRACT:

This paper introduces a system for real-time generation of digital surface models (DSM) based on an optical multi-camera system flown on board of a manned airplane or helicopter. The system consists of high end consumer cameras, GNSS/IMU system, and on-board computers for real-time data processing. Usually, generation of digital surface models from aerial imagery is done in an off-line process, leading to delayed availability of height data. The proposed system processes data in real time on board of the aircraft and downlinks the generated DSM to a ground station. This paper evaluates the GNSS/IMU on-line solution quality and its impact on dense stereo matching. The proposed real time sliding window based bundle adjustment significantly improves image orientations and DSM quality, allowing generation of detailed digital surface models with a resolution of 2*GSD. Experiments using two flight patterns are conducted over the city of Landsberg and the resulting DSMs are evaluated against a LiDAR generated reference point cloud. The online bundle adjustment is shown to minimize the effect of systematic GNSS/IMU offsets while adding only a limited delay.

1. INTRODUCTION

Generation of Digital Elevation Models (DEM) from airborne imagery is a one of the basic photogrammetric tasks, and the basis for many applications. Real-time processing expands the range of applications, in particular for crisis situations like natural disasters when up to date information is required. Based on real-time generated DEMs, damages on buildings or landslides can be derived quickly.

The conventional DSM generation process is divided into image acquisition, for example using UAVs, planes, helicopters or satellites, where images are recorded using optical camera systems. Later, these images are transferred via a data link or hard drives to a processing center, where the image and sensor data is processed using several steps, mainly consisting of image orientation, dense image matching and post-processing. Depending on the sensor platform and the availability of processing hard/software as well as skilled operators, the derivation of DEMs can take a long time.

In this work we present a workflow that is capable of generating digital elevation models in real time, synchronized to the image capture, and outputting a steady stream of height data with a constant latency. This work evaluates the influence of incremental bundle adjustment on DSM generation, which is performed directly after image acquisition using the measurements of an installed GNSS/IMU system (see section 2.1). Real-time DSM generation on conventinal CPUs is described in section 2.2. The proposed method was validated against LiDAR ground truth data on two image sets acquired with the 4K sensor system on a helicopter (see section 3).

2. METHOD

In this section, the real-time improvement of image orientations and the real-time DEM generation are described. The algorithms are tailored for optical sensor systems installed on aircrafts or helicopters in combination with a real-time capable GNSS/IMU. Experiments are conducted with the 4K sensor system, described in section 3.1.

2.1 Image Orientation

The improvement of the image orientations is necessary, as the real-time measured image attitudes are not accurate enough at pixel level to apply SGM based dense image matching. If the quality of the on-line GNSS/IMU data and the camera calibration is high enough, no further image orientation would be required.

Based on the specifications of most real-time capable GNSS/IMU systems, an on-line image based orientation is required for accurate and reliable operation of the real-time DSM matching. A complete block adjustment can only be performed after all images have been acquired, which conflicts with our goal of a low latency, real time system. Therefore, we use an incremental on-line bundle adjustment where image orientations are frozen after N images have been acquired. Thus our system can operate with a latency of $(N + 1) * t_e$, where t_e is the time between subsequent camera exposures. The idea is similar to (Engels et al., 2006), except that we use fewer free frames and consider GNSS/IMU measurements and a coarse reference DSM as additional observations.

Once a new image becomes available, the GNSS/IMU derived orientations and a coarse worldwide SRTM DEM (Rodriguez et al., 2005) are used to estimate the image footprint and to coarsely rectify the image using an affine transformation. BRISK (Leutenegger et al., 2011) features are extracted on these images and matched against features from previously acquired overlapping images. After RANSAC based outlier removal, the new pairwise matches are chained to multi-ray tie points. This strategy allows loop closing and connecting

of neighboring flight strips. Then a bundle block adjustment estimates the improved exterior orientation parameters for the last N active images. During the adjustment, tie point re-projection errors for all points connecting to the active images, deviations from GNSS/IMU measurements as well as height differences of the tie points to the coarse DEM are considered. Afterwards, the oldest active image pair is frozen, and used for dense matching. The adjustment part of our orientation estimation uses the Levenberg-Marquardt algorithm as implemented by the Ceres Solver package (Agarwal et al., 2018).

For multi-camera systems, such as the 4K system, image orientation needs to jointly consider images of all cameras, thus adjustment is performed using images from all cameras in a single estimation process.

2.2 DSM Generation

Once a new image with final orientation becomes available, it is matched to the previous image of the same camera. After epipolar rectification of the image pair, the fast three way SGBM implementation of OpenCV (Bradski, 2000) is used to estimate a dense disparity map. The three way SGBM only aggregates from left to right, top to down and right to left, instead of using the full 8 or 16 directions as the original SGM (Hirschmüller, 2008) algorithm, but it allows CPU and memory efficient computation at the price of reduced disparity quality.

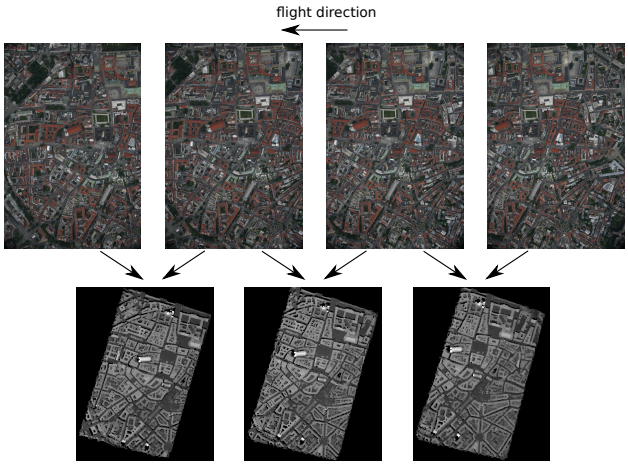

Figure 1. Pairwise dense matching and DSM generation.

Disparity maps are re-projected and re-sampled into the target coordinate system, creating one DSM for each image pair, as illustrated in figure 1. The DSM has lower resolution and less data volume than the original image, and can be downlinked to the ground station for further mosaicing into a complete DSM. When using images with a high forward and sideway overlap, most matching errors are removed by median based DSM merging.

3. EXPERIMENTAL EVALUATION

3.1 Hardware System

The whole work has been developed based on imagery acquired with the 4K camera systems developed at DLR (Kurz et al., 2012, Kurz et al., 2014). The 4K system can be mounted on the fuselage of a BO 105 or EC 135 helicopters as illustrated in figure 2.

Figure 2. Helicopter EC 135 with 4K camera sensor unit.

The general architecture of this system consists of a GNSS/IMU system for providing real-time orientation, three high end consumer DSLR cameras, standard computers for online-processing and a microwave data link for down-linking of data produced on board of the aircraft to a ground station for further dissemination.

Since each camera is connected to a dedicated PC via USB cables, acquired images can be immediately transferred to the PC for further processing. A dedicated operator PC or laptop is used to control the system during flight.

The cameras are triggered on a regular interval, and image and GNSS/IMU data is transferred to each camera PC. A software module reads images from the camera and associated GNSS/IMU data and uses them together with a known boresight angle and interior camera orientation to provide images with complete direct orientation. The bundle adjustment module is run on the operator PC and reads images and orientation data from the camera PCs. Refined orientation parameters are written back to the camera PCs. The DSM generation module running on each camera PC matches the last two images and provides the pairwise DSMs to the downlink module for transmission to the ground station. A DSM merging module running in the ground station performs tile based merging of the pairwise DSMs. Each tile that does not receive new DSM data within a certain amount of time is merged and send to further downstream processes.

3.2 Data sets

The described procedure was tested with imagery acquired over the city of *Landsberg am Lech* in Germany on 02-10-2018. The test area of $4 \times 4km$ size was acquired twice, first with linear flight strips and then with curved flight strips.

The data set with seven linear flight strips contains 788 images, 394 of them left resp. right looking. The strips were flown with a overlap of around 50% at a height of 500 meters above ground, using EOS-1Dx cameras equipped with 50 mm lenses, leading to a ground resolution of approximately 7 cm.

The data set with eight curved flight strips contains 658 images, 329 from each camera. The radius of one curved flight strip is around $1.5km$, the other flight parameter like flight height and overlap are similar to the data set with linear flight strips.

The real-time system was designed to be operated at a flight height of 1000 meters above ground, but due to clouds, only a flight height of 500 meters above ground was possible. As dense matching requires a high forward overlap of 75%, the image acquisition rate was increased from $1/3$ to $2/3$ Hz.

All measured projection centers are mapped in figure 3. Blue dots represent the projection centers of the linear flight strips, red dots the curved ones. Each point represents two images, one left- and one right-looking image.

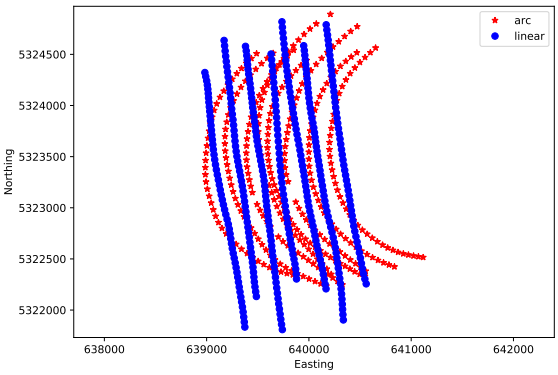

Figure 3. Curved and linear flight strips over *Landsberg am Lech*.

3.3 Real-time GNSS/IMU

On the 4K sensor system, two different real-time capable GNSS/IMU systems are integrated, the IGI-IId on EC-135 and BO-105 with fiber optic gyros operating at $128Hz$ or IGI-CompactMEMS on EC-135 working electro-mechanically with $400Hz$. Both GNSS/IMU systems have systems using differential GNSS with corrections from Omnistar XP resp. TerraStar. In general, highest accuracy in position and attitude can be reached in post-processing, whereas the performance of real-time solutions depend on many factors, which are also related to the direct situation in flight e.g. vibrations, flight speed, time for initialization etc. Crucial is the dynamical initial alignment in particular for the yaw angle, which is in our case important after GNSS outages caused by turning manoeuvres of the helicopter or by unfavourable flight patterns. The lower real-time accuracy is illustrated in figure 4, where the differences of the real-time vs. postprocessed solutions for the position and attitudes are plotted.

The differences in position and attitude for the curved flight lines (see figure 4 **a)** and **b)**) are higher than for the straight flight lines (see figure 4 **c)** and **d)**). Larger differences can be seen especially in the yaw angles at the beginning of the curved lines. Here, the initialization of the yaw angle failed due to the unfavorably curved flight lines.

Dense matching requires good relative image orientations, epipolar images with an vertical error of less than 0.5 pixel are required for good performance (Hirschmüller , Gehrig, 2009). When using images downsampled with a factor 2, orientation errors should be below 1 pixel in the original imagery, corresponding to an pitch error of 0.01° for the setup used in this paper. However, we found for the helicopter system that the real-time measured positions and attitudes of the GNSS/IMU system shows higher deviations, compromising the stereo matching quality. In particular, random orientation errors between subsequent images are higher than required.

3.4 DSM Evaluation

DSMs were generated using four different image orientation methods:

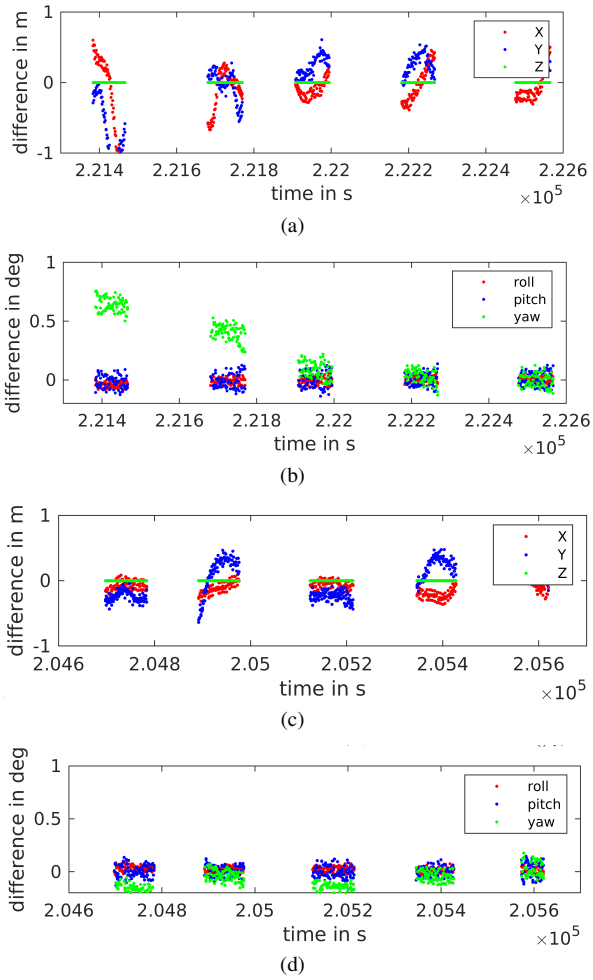

Figure 4. Differences between real-time and post-processed solutions for five curved flight lines **a)** and **b)** as well as for five straight flight lines **c)** and **d)** in the test area.

1. *Direct*: Use of on-line GPS/IMU data together with pre-defined boresight and interior camera calibration.

2. *Incremental* On line incremental bundle adjustment with $N = 7$ free images.

3. *2 strip*: bundle adjustment of the first two strips, followed by on-line incremental bundle adjustment with $N = 7$ free images,

4. *Postproc*: Offline processing: GNSS/IMU solution postprocessed using SAPOS, and high quality off-line bundle block adjustment.

We expect that the *2 strip* method is more reliable, as it can use image tie points between the first two strips to correct for GNSS/IMU offsets observed in Section 3.3. After the first two flight lines have been bundle adjusted, all but the last N images are fixed and pairwise DSMs generation is started.

The 3 way SGM dense matching is performed on images downscaled by a factor of two, resuting in DSMs with a grid spacing of 0.2 m. The pairwise DSM generation takes between 1 and 3 seconds, depending on the flight height, image overlap and height differences in the scene. As we had to increase the image acquisition frequency due to the lower flight heights,

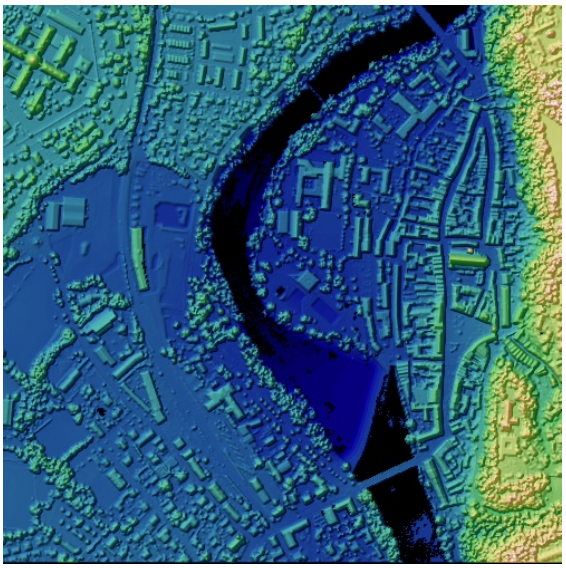

Figure 5. LiDAR Reference DSM used during evaluation.

the processing is not completely real-time, as a processing a stereo pair sometimes took longer than the 1.5 seconds between subsequent image acquisitions. For the originally planned flight at 1000 meters, processing would have been faster due to a reduced disparity range, and the time between image acquisition would have been 3 seconds, resulting a real time operation of the system.

Once all pairwise DSMs have been generated, they are merged by taking the median height for each DSM pixel and compared to a first pulse LiDAR point cloud obtained from the "Bayrisches Landesamt für Digitalisierung, Breitband und Vermessung". The LiDAR data was acquired in 2018, Fig. 5 shows the area used for evaluation.

Fig. 6 shows a close up of the different DSMs obtained from the linear flight strips. It is visible that the direct orientation yields a very noisy DSM, whereas the *Incremental* method with 7 free images significantly reduces noise, outliers and systematic offsets between stereo pairs. The *2 strip* method further reduces noise and add some finer details, at the expense of additional latency, since DSM generation only starts after the first two flight lines have been completed.

Height differences between the real-time DSM and the LiDAR DSM are shown in Fig. 7. Strong differences in deciduous vegetation and water areas are mostly due to acquisition time differences and different measurement principles. Additionally, remaining local rotations of the real-time DSM are visible, as indicated by systematic differences around buildings. Thus the larger errors in yaw could be reduced, but were not completely compensated by the incremental block adjustment.

Height difference statistics were computed between the real time DSMs and a LiDAR point cloud excluding water, vegetation and ground points beneath vegetation. The values reported in 1 confirm the results visualized in Fig. 6. Without incremental orientation, very high systematic and relative differences are reported, which are reduced when using online orientation. As expected, the *2 strip* approach performs better than the simpler *Incremental* method.

As expected from the orientation differences observed in Section 3.3, the linear flight strips show a higher accuracy than

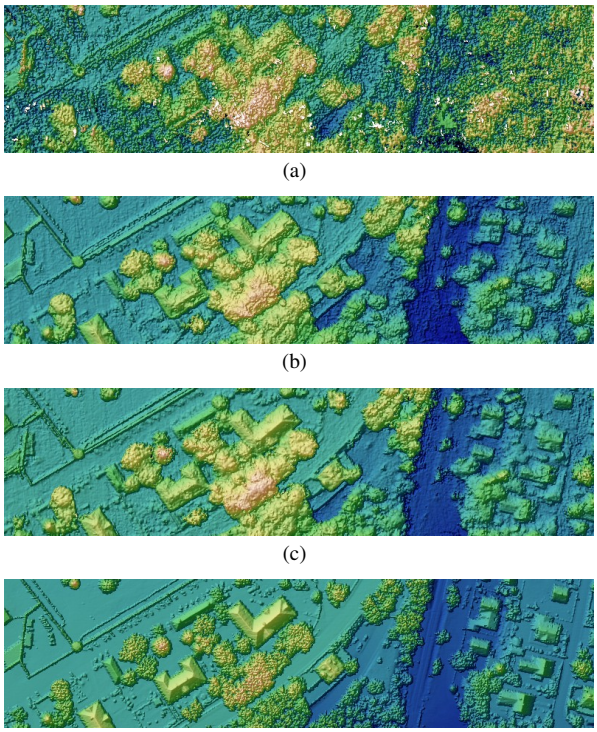

Figure 6. Visual comparison of different DSM generated using linear flight strips. (a) online direct georeferencing, (b) online bundle adjustment with 7 free images. (c) online bundle adjustment starting after two flight strips. (d) LiDAR ground truth

.

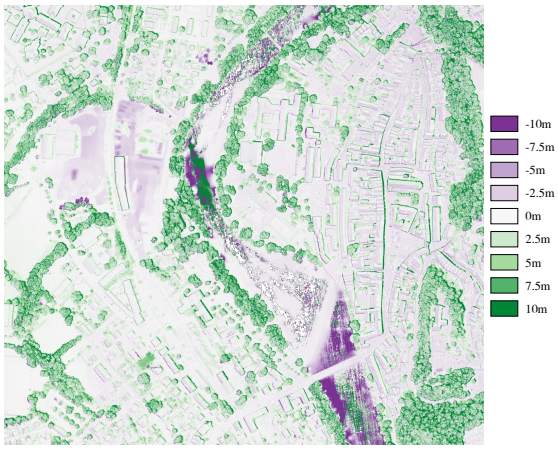

Figure 7. Height difference map between real-time DSM generated using the the 2 strip adjustment method and LiDAR point cloud. Systematic differences due to vegetation change are visible. Larger height differences around buildings indicate local rotations or shifts of the merged real-time DSM.

the curved strips, even when comparing the results of Line - Incremental with Curve - 2 strip. This is likely due to the strong yaw offsets during the first two flight strips that can not be recovered by the incremental bundle block adjustment later on.

The offline *Postproc* method represents the upper limit of the real-time DSM matching method. It performs significantly better than the real time version, as indicated by the much better MAD and good_1 values. Nevertheless, the incremental and two strip methods allow the real-time generation of digital surface modes with small absoute vertical offsets and relative

Experiment	RMSE	Median	MAD	good 1
Line - *Direct*	5.83	0.87	2.28	25.35 %
Line - *Incremental*	3.32	-0.44	0.82	51.89 %
Line - *2 strip*	3.00	-0.31	0.59	63.52 %
Line - *Postproc*	2.74	-0.39	0.26	77.31 %
Curve - *Direct*	8.41	0.05	3.26	19.26 %
Curve - *Incremental*	6.38	0.87	2.11	27.61 %
Curve - *2 strip*	5.33	0.46	1.78	30.94 %

Table 1. Height difference statistics between different real-time DSMs and LiDAR point cloud in meters. Median reports report absolute height error, median absolute deviation (MAD) report relative differences. RMSE and good 1 (percentage of DSM pixels within 1m of the LiDAR point cloud) indicate the overall accuracy.

height accuracy better than 1 meter. Depending on the application, vertical shifts with respect to the reference data are tolerable as they can be compensated by DSM alignment when performing tasks such as detection of building or landscape changes.

4. CONCLUSION

We describe a system for real time DSM generation using the helicopter based 4K camera system. The real-time solution of the GNSS/IMU system of the 4K System as flown on the BO 105 helicopter is not accurate enough for for dense matching. In particular, higher errors in the yaw angle and reduced GNSS availability during turning manoeuvres, lead to errors in epipolar geometry, resulting in sparse and noisy DSMs with local deformations and height errors.

When using incremental bundle adjustment, dense matching in real-time is possible and yields DSMs with consistent heights, and a resolution $2 * GSD$ of the input imagery. Comparison against LiDAR point clouds show a usable height accuracy, but the errors in yaw angle could not be compensated completely.

Best results were obtained when performing a full adjustment of the first two strips, allowing for better compensation of IMU uncertainty by the GNSS positions and image based measurement. Future work will include further development of the incremental adjustment algorithm, especially improving the tie point matching by use of GNSS/IMU a-priori knowledge, and improved consideration of GNSS/IMU uncertainty during block adjustment.

REFERENCES

Agarwal, Sameer, Mierle, Keir, Others, 2018. Ceres solver. http://ceres-solver.org.

Bradski, G., 2000. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*.

Engels, Chris, Stewénius, Henrik, Nistér, David, 2006. Bundle adjustment rules. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVI - 3, 266–271.

Hirschmüller, H., Gehrig, S., 2009. Stereo matching in the presence of sub-pixel calibration errors. *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 437–444.

Hirschmüller, Heiko, 2008. Stereo Processing by Semi-Global Matching and Mutual Information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30, 328 - 341. http://elib.dlr.de/55367.

Kurz, F., Rosenbaum, D., Meynberg, O., Mattyus, G., Reinartz, P., 2014. Performance of a real-time sensor and processing system on a helicopter. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-1, 189–193.

Kurz, F., Türmer, S., Meynberg, O., Rosenbaum, D., Runge, H., Reinartz, P., Leitloff, J., 2012. Low-cost Systems for real-time Mapping Applications. *Photogrammetrie Fernerkundung Geoinformation*, 159–176.

Leutenegger, S., Chli, M., Siegwart, R. Y., 2011. Brisk: Binary robust invariant scalable keypoints. *2011 International Conference on Computer Vision*, 2548–2555.

Rodriguez, E., Morris, C.S., Belz, J.E., Chapin, E.C., Martin, J.M., Daffer, W., Hensley, S., 2005. An assessment of the SRTM topographic products. Technical Report Technical Report JPL D-31639, Jet Propulsion Laboratory, Pasadena, California.